

0570
1221

#8

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/887,784

DATE: 01/02/2002
TIME: 13:13:40

Input Set : A:\3759-0115P.ST25.txt
Output Set: N:\CRF3\01022002\I887784.raw

ENTERED

3 <110> APPLICANT: BJORN, Sara et al
 5 <120> TITLE OF INVENTION: NOVEL FLUORESCENT PROTEINS
 7 <130> FILE REFERENCE: 3759-0115P
 9 <140> CURRENT APPLICATION NUMBER: US 09/887,784
 10 <141> CURRENT FILING DATE: 2001-06-19
 12 <160> NUMBER OF SEQ ID NOS: 24
 14 <170> SOFTWARE: PatentIn version 3.0
 17 <210> SEQ ID NO: 1
 18 <211> LENGTH: 720
 19 <212> TYPE: DNA
 20 <213> ORGANISM: Aequoria Victoria
 22 <220> FEATURE:
 23 <221> NAME/KEY: CDS
 24 <222> LOCATION: (1)...(717)
 26 <400> SEQUENCE: 1
 27 atg gtg agc aag ggc gag gag ctg ttc acc ggg gtg gtg ccc atc ctg 48
 28 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 29 1 5 10 15
 31 gtc gag ctg gac ggc gac gta aac ggc cac aag ttc agc gtg tcc ggc 96
 32 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
 33 20 25 30
 35 gag ggc gag ggc gat gcc acc tac ggc aag ctg acc ctg aag ttc atc 144
 36 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
 37 35 40 45
 39 tgc acc acc ggc aag ctg ccc gtg ccc tgg ccc aca cta gtg acc acc 192
 40 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
 41 50 55 60
 43 ctg tct tac ggc gtg cag tgc ttc agc cgc tac ccc gac cac atg aag 240
 44 Leu Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
 45 65 70 75 80
 47 cag cac gac ttc aag tcc gcc atg ccc gaa ggc tac gtc cag gag 288
 48 Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
 49 85 90 95
 51 cgc acc atc ttc aag gac gac ggc aac tac aag acc cgc gcc gag 336
 52 Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
 53 100 105 110
 55 gtg aag ttc gag ggc gac acc ctg gtg aac cgc atc gag ctg aag ggc 384
 56 Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
 57 115 120 125
 59 atc gac ttc aag gag gac ggc aac atc ctg ggg cac aag ctg gag tac 432
 60 Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
 61 130 135 140
 63 aac tac aac agc cac aac gtc tat atc atg gcc gac aag cag aag aac 480
 64 Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
 65 145 150 155 160
 67 ggc atc aag gtg aac ttc aag atc cgc cac aac atc gag gac ggc agc 528
 68 Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/887,784

DATE: 01/02/2002
TIME: 13:13:40

Input Set : A:\3759-0115P.ST25.txt
Output Set: N:\CRF3\01022002\I887784.raw

69	165	170	175	
71	gtg cag ctc gcc gac cac tac cag cag aac acc ccc atc ggc gac ggc			576
72	Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly			
73	180	185	190	
75	ccc gtg ctg ctg ccc gac aac cac tac ctg agc acc cag tcc gcc ctg			624
76	Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu			
77	195	200	205	
79	agc aaa gac ccc aac gag aag cgc gat cac atg gtc ctg ctg gag ttc			672
80	Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe			
81	210	215	220	
83	gtg acc gcc gcc ggg atc act ctc ggc atg gac gag ctg tac aag			717
84	Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys			
85	225	230	235	
87	taa			720
89	<210> SEQ ID NO: 2			
90	<211> LENGTH: 239			
91	<212> TYPE: PRT			
92	<213> ORGANISM: Aequoria Victoria			
94	<400> SEQUENCE: 2			
95	Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu			
96	1	5	10	15
97	Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly			
98	20	25	30	
99	Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile			
100	35	40	45	
101	Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr			
102	50	55	60	
103	Leu Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys			
104	65	70	75	80
105	Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu			
106	85	90	95	
107	Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu			
108	100	105	110	
109	Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly			
110	115	120	125	
111	Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr			
112	130	135	140	
113	Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn			
114	145	150	155	160
115	Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser			
116	165	170	175	
117	Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly			
118	180	185	190	
119	Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu			
120	195	200	205	
121	Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe			
122	210	215	220	
123	Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys			
124	225	230	235	

1 Aug 2002
RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/887,784

DATE: 01/02/2002
TIME: 13:13:40

Input Set : A:\3759-0115P.ST25.txt
Output Set: N:\CRF3\01022002\I887784.raw

126 <210> SEQ ID NO: 3
127 <211> LENGTH: 720
128 <212> TYPE: DNA
129 <213> ORGANISM: Aequoria Victoria
131 <220> FEATURE:
132 <221> NAME/KEY: CDS
133 <222> LOCATION: (1)...(717)
135 <400> SEQUENCE: 3
136 atg gtg agc aag ggc gag gag ctg ttc acc ggg gtg gtg ccc atc ctg 48
137 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
138 1 5 10 15
140 gtc gag ctg gac ggc gac gta aac ggc cac aag ttc agc gtg tcc ggc 96
141 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
142 20 25 30
144 gag ggc gag ggc gat gcc acc tac ggc aag ctg acc ctg aag ttc atc 144
145 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
146 35 40 45
148 tgc acc acc ggc aag ctg ccc gtg ccc tgg ccc aca cta gtg acc acc 192
149 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
150 50 55 60
152 ctg tct tac ggc gtg cag tgc ttc agc cgc tac ccc gac cac atg aag 240
153 Leu Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
154 65 70 75 80
156 cag cac gac ttc ttc aag tcc gcc atg ccc gaa ggc tac gtc cag gag 288
157 Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
158 85 90 95
160 cgc acc atc ttc ttc aag gac gac ggc aac tac aag acc cgc gcc gag 336
161 Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
162 100 105 110
164 gtg aag ttc gag ggc gac acc ctg gtg aac cgc atc gag ctg aag ggc 384
165 Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
166 115 120 125
168 atc gac ttc aag gag gac ggc aac atc ctg ggg cac aag ctg gag tac 432
169 Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
170 130 135 140
172 aac tac aac agc cac aac gtc tat atc atg gcc gac aag cag aag aac 480
173 Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
174 145 150 155 160
176 ggc atc aag gtg aac ttc aag atc cgc cac aac atc gag gac ggc agc 528
177 Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
178 165 170 175
180 gtg cag ctc gcc gac cac tac cag cag aac acc ccc atc ggc gac ggc 576
181 Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
182 180 185 190
184 ccc gtg ctg ctg ccc gac aac cac tac ctg agc acc cag tcc gcc ctg 624
185 Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
186 195 200 205
188 agc aaa gac ccc aac gag aag cgc gat cac atg gtc ctc cta ggg ttc 672
189 Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Gly Phe

148-700
RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/887,784

DATE: 01/02/2002
TIME: 13:13:40

Input Set : A:\3759-0115P.ST25.txt
Output Set: N:\CRF3\01022002\I887784.raw

190 210 215 220
192 gtg acc gcc gcc ggg atc act ctc ggc atg gac gag ctg tac aag 717
193 Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
194 225 230 235
196 taa 720
198 <210> SEQ ID NO: 4
199 <211> LENGTH: 239
200 <212> TYPE: PRT
201 <213> ORGANISM: Aequoria Victoria
203 <400> SEQUENCE: 4
204 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
205 1 5 10 15
206 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
207 20 25 30
208 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
209 35 40 45
210 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
211 50 55 60
212 Leu Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
213 65 70 75 80
214 Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
215 85 90 95
216 Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
217 100 105 110
218 Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
219 115 120 125
220 Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
221 130 135 140
222 Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
223 145 150 155 160
224 Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
225 165 170 175
226 Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
227 180 185 190
228 Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
229 195 200 205
230 Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Gly Phe
231 210 215 220
232 Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
233 225 230 235
235 <210> SEQ ID NO: 5
236 <211> LENGTH: 717
237 <212> TYPE: DNA
238 <213> ORGANISM: Aequoria Victoria
240 <220> FEATURE:
241 <221> NAME/KEY: CDS
242 <222> LOCATION: (1)...(714)
244 <400> SEQUENCE: 5
245 atg agt aaa gga gaa gaa ctt ttc act gga gtt gtc cca att ctt gtt 48

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/887,784

DATE: 01/02/2002
TIME: 13:13:40

Input Set : A:\3759-0115P.ST25.txt
Output Set: N:\CRF3\01022002\I887784.raw

246	Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val				
247	1	5	10	15	
249	gaa tta gat ggc gat gtt aat ggg caa aaa ttc tct gtt agt gga gag				96
250	Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser Gly Glu				
251	20	25	30		
253	ggt gaa ggt gat gca aca tac gga aaa ctt acc ctt aaa ttt att tgc				144
254	Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys				
255	35	40	45		
257	act act ggg aag cta cct gtt cca tgg cca acg ctt gtc act act ctc				192
258	Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu				
259	50	55	60		
261	tct tat ggt gtt caa tgc ttt tct aga tac cca gat cat atg aaa cag				240
262	Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln				
263	65	70	75	80	
265	cat gac ttt ttc aag agt gcc atg ccc gaa ggt tat gta cag gaa aga				288
266	His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg				
267	85	90	95		
269	act ata ttt tac aaa gat gac ggg aac tac aag aca cgt gct gaa gtc				336
270	Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val				
271	100	105	110		
273	aag ttt gaa ggt gat acc ctt gtt aat aga atc gag tta aaa ggt att				384
274	Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile				
275	115	120	125		
277	gat ttt aaa gaa gat gga aac att ctt gga cac aaa atg gaa tac aat				432
278	Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu Tyr Asn				
279	130	135	140		
281	tat aac tca cat aat gta tac atc atg gca gac aaa cca aag aat ggc				480
282	Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Pro Lys Asn Gly				
283	145	150	155	160	
285	atc aaa gtt aac ttc aaa att aga cac aac att aaa gat gga agc gtt				528
286	Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp Gly Ser Val				
287	165	170	175		
289	caa tta gca gac cat tat caa caa aat act cca att ggc gat ggc cct				576
290	Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro				
291	180	185	190		
293	gtc ctt tta cca gac aac cat tac ctg tcc acg caa tct gcc ctt tcc				624
294	Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser				
295	195	200	205		
297	aaa gat ccc aac gaa aag aga gat cac atg atc ctt ctt gag ttt gta				672
298	Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu Phe Val				
299	210	215	220		
301	aca gct gct ggg att aca cat ggc atg gat gaa ggg tac aag				714
302	Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Gly Tyr Lys				
303	225	230	235		
305	taa				717
307	<210> SEQ ID NO: 6				
308	<211> LENGTH: 238				
309	<212> TYPE: PRT				
310	<213> ORGANISM: Aequoria Victoria				

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/887,784

DATE: 01/02/2002

TIME: 13:13:41

Input Set : A:\3759-0115P.ST25.txt

Output Set: N:\CRF3\01022002\I887784.raw